

PROGRAMA DEL CURSO

GESTIÓN DE DATOS EN LA NUBE

X€

40h

Incluye Certificado WAT

Curso integral: Big Data fundamentals, estadística, bases de datos, SQL avanzado, Python para datos, pipelines ETL, Google Cloud, machine learning aplicado y governance de datos.

PROGRAMA

UNIDAD 1: INTRODUCCIÓN A BIG DATA Y CONCEPTOS FUNDAMENTALES

Entiende características de Big Data (volumen, velocidad, variedad), desafíos de procesamiento distribuido, diferencia Big Data tradicional vs cloud-native y evalúa tecnologías. Aprende arquitectura de sistemas distribuidos.

UNIDAD 2: ESTADÍSTICA FUNDAMENTAL PARA ANALÍTICA

Domina estadística descriptiva, distribuciones de probabilidad, teorema central del límite, inferencia estadística y pruebas de hipótesis. Aplica conceptos para validar insights y evitar errores estadísticos comunes.

UNIDAD 3: BASES DE DATOS Y MODELADO DE DATOS

Comprende modelo relacional y diseño de esquemas ER. Aplica normalización (1NF, 2NF, 3NF), crea índices para performance, explora NoSQL (documentos, grafos, key-value) y polyglot persistence.

UNIDAD 4: SQL AVANZADO Y TRANSFORMACIÓN DE DATOS

Domina SQL avanzado: CTEs recursivas, window functions (ROW_NUMBER, RANK, LAG, LEAD), pivoting/unpivoting, subconsultas complejas y optimización de queries para performance.

UNIDAD 5: PYTHON PARA ANÁLISIS Y PROCESAMIENTO DE DATOS

Utiliza Python con Pandas, NumPy y Matplotlib para análisis exploratorio. Manipula DataFrames, realiza limpieza de datos, crea visualizaciones y integra Python con bases de datos.

UNIDAD 6: CICLO DE VIDA DE DATOS Y PIPELINES ETL

Diseña pipelines ETL robustos: extrae de múltiples fuentes, transforma datos, carga en data warehouse. Implementa validación, error handling, idempotencia y documentación de pipelines.

UNIDAD 7: GOOGLE CLOUD PLATFORM PARA GESTIÓN DE DATOS

Implementa soluciones GCP: BigQuery como data warehouse, Dataflow para pipelines, Cloud Storage para data lake, integraciones nativas. Configura infraestructura y autorización.

UNIDAD 8: MACHINE LEARNING PARA ANÁLISIS Y PREDICCIÓN

Aplica machine learning: conceptos de supervised/unsupervised learning, preparación de datos, feature engineering, crear modelos predictivos, evaluación y optimización. BQML para ML sin código complejo.

UNIDAD 9: GOVERNANCE, SEGURIDAD Y CALIDAD DE DATOS

Implementa governance: privacidad (GDPR, CCPA), control de acceso (RBAC), auditoría, compliance. Asegura calidad de datos: definición de métricas, validaciones, monitoreo y mejora continua.

HERRAMIENTAS

● Google Cloud Platform

● BigQuery

● Dataflow

● Python

● Firestore

QUÉ VAS A CONSEGUIR

- Explicar características de Big Data y evolución de tecnologías
- Aplicar estadística descriptiva e inferencia estadística
- Diseñar esquemas relacionales normalizados
- Escribir queries SQL avanzadas con CTEs y window functions
- Manipular datos con Pandas y NumPy
- Diseñar y implementar pipelines ETL idempotentes
- Crear data warehouse en BigQuery optimizado
- Implementar modelos predictivos con machine learning
- Definir e implementar governance y seguridad de datos

¿PARA QUIÉN ES ESTE CURSO?

Para data engineers, data scientists y profesionales que necesitan soluciones analíticas empresariales completas.

- Data engineers implementando data warehouses
- Data scientists aplicando ML a datos masivos
- Analistas que necesitan competencias full-stack
- Arquitectos diseñando soluciones Big Data
- Equipos de datos en transformación digital
- Profesionales buscando dominar datos de extremo a extremo

wearetech.es

We are tech | WAT

¿Preparado para dar el siguiente paso?

wearetech.es